Evaluation of a Method for Proactively Managing the Evolving Scope of a Software Product Line

Karina Villela, Jörg Dörr, and Isabel John

Fraunhofer Institute for Experimental Software Engineering
Fraunhofer-Platz 1, 67663 Kaiserslautern, Germany
{Karina.Villela,Joerg.Doerr,Isabel.John}@iese.fraunhofer.de

Abstract. **[Context and motivation]** PLEvo-Scoping is a method intended to help Product Line (PL) scoping teams anticipate emergent features and distinguish unstable from stable features, with the aim of preparing their PL for likely future adaptation needs. **[Question/problem]** This paper describes a quasi-experiment performed to characterize PLEvo-Scoping in terms of adequacy and feasibility. **[Principal ideas/results]** This quasi-experiment was performed by two scoping teams in charge of scoping the same PL, where one scoping team applied first an existing PL scoping approach and then PLEvo-Scoping, while the other scoping team interweaved activities from both. The two approaches achieved similar results: The method could be applied in just one day, and it was considered adequate and feasible. **[Contribution]** Ideas on how to improve the method and its tool support have been obtained, and similar results are expected from other professionals facing the problem of evolution-centered PL scoping. However, further empirical studies should be performed.

Keywords: Quasi-experiment, Empirical Study, Product Line Scoping, Requirements Volatility Analysis, Product Line Evolution Planning.

1 Introduction

Product Line (PL) Engineering is a software development approach that aims at exploiting commonalities and predicted variabilities among software products that strongly overlap in terms of functionality [1,2]. According to Knauber and Succi [3], PLs are already intended to capture the evolution of software products and to last for a fairly long time. In this context, one of the most important aspects to consider is the ability of PLs themselves to evolve and change [3]. However, Savolainen and Kuusela [4] emphasize that any given design can only handle a limited number of different kinds of changes and, therefore, it is crucial to predict what kind of changes will be required during the lifespan of a PL.

Current PL engineering methods [1,2,5] address pre-planned and more straightforward proactive changes across products or different versions of products. They do not support the prediction of the not-so-straightforward future changes in products and features, which are often triggered by change requests from inside or outside the organization (such as a change due to the decision of a technology provider to

R. Wieringa and A. Persson (Eds.): REFSQ 2010, LNCS 6182, pp. 113–127, 2010.

discontinue the production of a hardware component). Therefore, a good commonality and variability analysis is not enough [6].

PL scoping is the PL requirements engineering step in charge of deciding in which parts of an organization's products systematic reuse is economically useful and should thus be supported by a PL infrastructure [7,8]. When focusing on this phase, it is customary to use the identified requirements that are outside the current scope as candidates for future addition as well as to use those requirements to evaluate the design of the current PL. However, this does not go as far as an explicit volatility analysis.

The method PLEvo-Scoping (Product Line Evolution Support at Scoping) was defined to complement and extend PL scoping approaches by helping the PL scoping team anticipate emergent features and distinguish unstable from stable PL features. The aim is to prepare the PL for likely future adaptation needs by planning for changes beforehand. Thus, the PL scoping approach can keep its focus on current, planned, and potential PL products and features, while PLEvo-Scoping addresses the prediction and planning for the evolution of these products and features, while also contributing to the discovery of further innovative features.

This article describes a quasi-experiment performed to characterize the adequacy and feasibility of PLEvo-Scoping with a twofold purpose: 1) obtaining feedback from PL practitioners on how to improve the method; and 2) providing first empirical data to help PL organizations decide about trying out the method. Two scoping teams (each composed of three professionals) had the task of scoping the same PL by using PLEvo-Scoping together with an existing PL scoping approach, but with two different arrangements of sub-tasks. We believe this quasi-experiment has been a relevant step towards the evaluation of the method against other potential approaches, because any PL organization is only willing to test a new method if there is initial evidence that the benefits to be obtained are worth the effort required to apply it. It is not the goal of this paper to describe the state of the art related to PLEvo-Scoping or the method itself in detail; this information can be found in [9] and [10].

In the reported quasi-experiment, PLEvo-Scoping was applied in the Ambient Assisted Living (AAL) domain for the second time. Products in this domain support elderly people in staying at home longer instead of moving to a nursing home. However, the two applications of the method in this domain were completely different. The first one [9] can be considered an assertion study [11], in which one of the authors applied the method herself after studying the application domain, while domain experts just validated the method results, indicating the necessary corrections. The object was the specific AAL services to be provided by an AAL demonstrator. Concerning the quasi-experiment, the method was exclusively applied by the two PL scoping teams and the object was the AAL platform PL. An AAL platform is an architectural component that provides infrastructure services on top of which the specific AAL services are provided.

In the next sections, we provide an overview of PLEvo-Scoping (Section 2) and briefly discuss how it interfaces with existing PL scoping approaches (Section 3), so that the quasi-experiment can be understood better. Section 4 reports the quasi-experiment and is organized as follows: definition, planning, operation, data analysis, and comments and interpretation of the results. The last section contains our conclusions and the next stages of this research.

2 Method Overview

PLEvo-Scoping consists of four steps to be carried out by the PL scoping team, which is generally composed of people with the following roles [8]: scoping expert, PL manager, and domain expert, the latter with either the technical or the market point of view.

The first step is *Preparation for Volatility Analysis*, which establishes the basis for the volatility analysis and is made up of the following activities:

- *Activity 1:* Establish the timeframe that restricts the current volatility analysis, and
- *Activity 2:* Identify/update the types of system components that are generally involved in the assembly of the planned PL products.

The second step is called *Environment Change Anticipation* and has the purpose of identifying and characterizing facts that may take place in the PL's environment within the pre-established timeframe, and that may allow or require adaptations in the PL. This step comprises the following activities:

- *Activity 3:* Identify the actors that play a role in the PL's environment and who give rise to or realize facts that may affect the PL,
- *Activity 4:* Identify and characterize facts that may be caused or realized by the identified actors and have the potential for changing the PL's environment,
- *Activity 5:* Verify the perspective of new actors playing a part in the PL's environment within the volatility analysis timeframe and characterize how these actors may change such an environment, and
- *Activity 6:* Classify the previously characterized facts according to their relevance, in order to decide whether and when they should have their impact in terms of adaptation needs analyzed.

The next step is called *Change Impact Analysis*. Its purpose is to analyze the impact of the identified facts on the PL and consists of:

- *Activity 7:* Identify the adaptation needs that may be allowed or required in the PL as a consequence of the previously identified facts,
- *Activity 8:* Characterize the adaptation needs by identifying the PL features to be affected by them, and by estimating their business impact, technical penalty, and technical risk, and
- *Activity 9:* Classify the adaptation needs according to relevance, in order to decide whether and when the inclusion of an adaptation need should be planned.

Once the most relevant adaptation needs have been selected, it is time for *PL Evolution Planning*. The idea is to establish when and how relevant adaptation needs are expected to be introduced into the PL, and prepare it for accommodating the adaptation needs beforehand. The activities that make up this step are:

- *Activity 10:* Determine when and in which products relevant adaptations are expected to be introduced, which gives rise to the *PL Evolution Map*,
- *Activity 11:* Analyze the alternative solutions for dealing with relevant adaptation needs, in terms of effort, cost, effectiveness, and strategic alignment,
- *Activity 12:* Select the best alternatives for dealing with the adaptation needs, and
- *Activity 13:* Revise the *PL Evolution Map* in order to adjust it to the alternative solutions selected, if necessary.

Procedural descriptions, checklists, guidelines, and templates of documents are provided to support the PL scoping team in carrying out each activity of the method. In addition, optional activities have been identified in order to address the different levels of experience in the domain that different teams may have, and also to take into account time restrictions, by providing a light-weight variant of the method. Most optional activities use systematic reasoning to identify and characterize facts and adaptation needs that otherwise would have to be identified from scratch. Therefore, they are very important in terms of the completeness of the method results, especially when the PL scoping team is not very experienced in the application domain. The other optional activities aim at making the method scalable when too many facts or adaptation needs are identified.

After applying the method, the most relevant adaptation needs likely to be allowed or required within the volatility analysis timeframe are expected to be identified and included in a plan for PL evolution. Excerpts from output documents can be found in [9] and [10].

3 Integration into the Scoping Process

PLEvo-Scoping is expected to be carried out after or in parallel to the standard PL scoping process, provided the mandatory method inputs are available (*Description of Products* and *List of PL Features*). Broadly speaking, the method takes those mandatory inputs together with relevant information for reasoning on requirements volatility, and helps to identify a set of adaptations likely to be requested or allowed in an established future timeframe, pointing out on a map how the PL features will be affected by these adaptations. New features may arise; others may become obsolete; it may be revealed that a feature will be affected by many potential future adaptations and therefore is an unstable feature; or it may be revealed that an adaptation will affect many different features and thus requires special attention. Therefore, the *List of PL Features* can be improved, by including the newly identified features, characterizing the PL features according to their volatility, and changing the priority of some PL features based on their volatility or relevance for a certain adaptation need. The *Product Release Plan*, which is not a mandatory but a much desired method input, is augmented with the indication of when and in which products relevant adaptation needs are expected to be introduced. Furthermore, the analysis of alternative solutions for dealing with the adaptation needs may affect the prioritization of assets to be built for the PL and will indicate what kind of evolvability should be built into them.

PuLSE-Eco [7,8] is the technical component of the PuLSE® methodology [5] in charge of supporting PL scoping. This technical component, which is the best known and best documented scoping approach so far, is used below to illustrate the integration between PLEvo-Scoping and a concrete PL scoping process, because this approach was used in the quasi-experiment that will be described in the next sections.

PuLSE-Eco's generic process comprises three phases: Product Line Mapping, Domain Potential Assessment, and Reuse Infrastructure Scoping (see the top part of Figure 1). During the *Product Line Mapping* phase, the products and features of a PL are identified and the distribution of features to products is established. The goal is to provide an overview of the PL. During the *Domain Potential Assessment* phase, sub-domains are analyzed according to assessment dimensions in order to support the decision on whether to include these sub-domains in the PL infrastructure or not. During the *Reuse Infrastructure Scoping* phase, existing and planned assets are identified and their reuse potential is quantified. The aim of this last phase is to plan the PL infrastructure and identify development needs. Figure 1 shows the main outputs of these phases and indicates in which PLEvo-Scoping activities they are requested (mandatory inputs) or can make a contribution (optional inputs). On the other hand, the outputs of PLEvo-Scoping are used to update the *(Quantified) Product Feature Matrix* and the *Product Release Plan* as already explained, once the *(Quantified) Product Feature Matrix* contains the *List of PL Features*; as well as to augment the *Domain Assessment Report* with additional information about the stability of the domains, one of the assessment dimensions proposed in PuLSE-Eco [7]. PLEvo-Scoping can also be integrated into other scoping processes, as they all provide similar results [7].

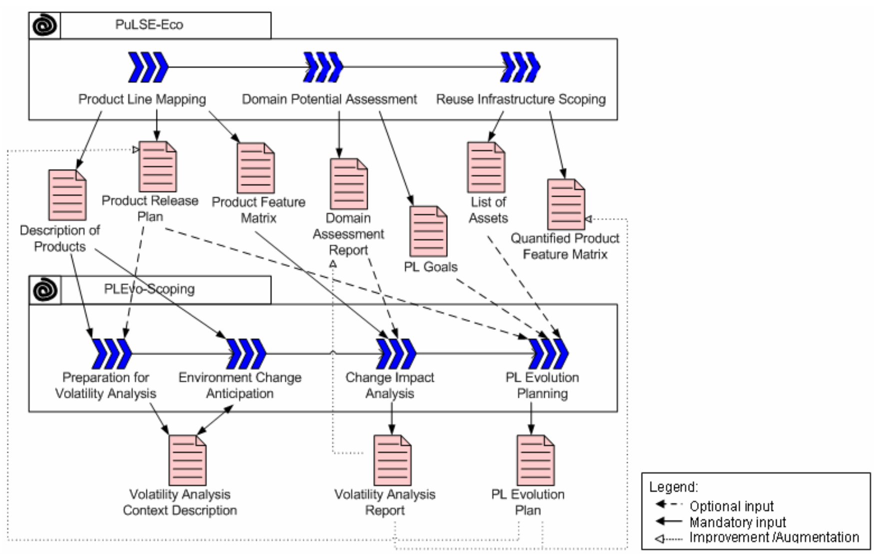

Fig. 1. Interface between a PL scoping process and PLEvo-Scoping

4 Quasi-experiment

4.1 Definition

The goal of this quasi-experiment was to characterize the adequacy and feasibility of PLEvo-Scoping, by collecting the perception of the quasi-experiment participants as well as some quantitative measures. These empirical data are expected to support PL organizations in making the decision to try out the method, which will give us the opportunity to perform further empirical studies, such as a case study in a software company. In addition, feedback provided by the quasi-experiment participants should be used to improve the method. According to the template proposed in the

Goal/Question/Metric method [12], the goal of this quasi-experiment can be represented as follows: *Analyze* the PLEvo-Scoping method *for the purpose of* characterization *with respect to* its adequacy and feasibility *from the point of view of* the researchers, PL managers, and domain experts *in the context of* Fraunhofer IESE's employees in charge of scoping an Ambient Assisted Living PL.

Concerning the context, Fraunhofer IESE has a research program on Ambient Assisted Living (AAL), which is funded by various sources in the form of independent projects. Since there are a lot of overlaps between the different projects and the AAL platform is of strategic interest for the institute, the research program leader needed scoping of an AAL platform PL, so that a PL infrastructure capable of supporting the reuse of the commonalities could be created. He also agreed to using PLEvo-Scoping to extend Fraunhofer's approach for PL scoping (PuLSE-Eco [7,8]), since a simplified version of the method had already been applied to anticipate the evolution of a single product in this domain and actually provided insightful ideas on features for the next version of the product [9]. Therefore, there was a real problem to be solved by professionals at Fraunhofer IESE, which, on the one side, strongly limited the degree of control over the quasi-experiment, but, on the other side, increased the possibility of generalizing the quasi-experiment results to other companies with PL competencies. The AAL Platform PL scoping was not carried out as part of a research project, but rather as an internal project of strategic relevance.

4.2 Planning

From the goal definition, two propositions were defined: P1) PLEvo-Scoping is adequate, and P2) PLEvo-Scoping is feasible. It was established for this quasi-experiment that PLEvo-Scoping would be considered adequate (P1) if its immediate benefits (in terms of changes in the outputs of the scoping process and information made available to guide further activities and decisions) were considered to support PL evolution, according to the judgment of the quasi-experiment participants; and feasible (P2) if the obtained benefits were worth the effort required for applying the method, which would also be judged by the quasi-experiment participants. In both cases, the quasi-experiment participants should be provided with the results of the scoping process and support their judgment with both qualitative and quantitative information.

By using the Goal/Question/Metric method [12], the quasi-experiment propositions were broken down into questions and metrics. Each metric was defined in terms of meaning, type of measure, measure scale, source, and collecting procedure [13]. The acceptance criteria were also defined during the planning of the quasi-experiment and will be presented in Subsection 4.4 (Data Analysis).

This quasi-experiment was designed to be performed in two days by two scoping teams in charge of separately scoping the AAL platform as a PL. Group 1 applied treatment 1, which consisted of first using PuLSE-Eco to conduct the scoping process, already taking into consideration the evolution concern, and then applying PLEvo-Scoping; while Group 2 applied treatment 2, which consisted of interweaving the activities of PuLSE-Eco and PLEvo-Scoping. The purpose of treatment 1 was to allow a clear distinction between the results before and after applying PLEvo-Scoping, while the purpose of treatment 2 was to avoid the confounding factor (present in treatment 1) of providing the scoping team with extra time to think about PL scope evolution after applying PuLSE-Eco. This confounding factor would put in doubt whether the obtained benefits were a result of this extra time or a result of the application of PLEvo-Scoping. In addition, the application of PLEvo-Scoping in the two treatments aimed at strengthening the validity of the results through corroboration. Two days were allocated to each treatment due to time restrictions.

Each group consisted of three people: one PL manager and two domain experts, one with the technical point of view and the other one with the market point of view. All participants were selected by the leader of the AAL research program, taking into account their profiles and involvement in AAL research projects. Two experts on the respective two methods (PuLSE-Eco and PLEvo-Scoping) were allocated to guide the pertinent part of the scoping process in both treatments. Neither the PuLSE-Eco nor the PLEvo-Scoping expert was involved in the AAL research program, so their role in the scoping process was comparable to that of an external consultant.

As PL scoping teams are not generally composed of many people, all necessary roles were represented, and the PL scoping process was to take just a couple of days [8]; this design was realistic compared to industrial settings.

The threats to the validity of this quasi-experiment have been analyzed based on the set of threats to validity provided in [14]. We have addressed most validity threats:

- *Fishing*: Subjective classifications and measures were only provided by the participants of the quasi-experiment; two types of open questions were included; two people with no special expectations in the quasi-experiment results were involved in its data analysis.
- *Reliability of measures*: The quasi-experiment participants defined subjective measures and/or provided values for them based on objective measures; the instruments were revised by three people with different profiles (one M.Sc. student, one PL professional, and one expert in empirical studies).
- *Mono-method bias*: Both quantitative and qualitative measures were used; measures were cross-checked whenever possible; one group's contributions were confirmed by the other group.
- *Interaction of selection and treatment*: One representative of each expected role was allocated to each group and all participants answered a profile questionnaire to check whether they were really appropriate representatives of the roles they were expected to have.

A discussion of further threats (*reliability of treatment implementation*, *diffusion or imitation of treatments, hypothesis guessing, compensatory rivalry, and resentful demoralization*) can be found in [13]. As is common in (quasi-)experiments, some validity threats had to be accepted:

- *Low statistical power*: The number of subjects in this quasi-experiment made it impossible to perform any statistical analysis.
- *Random heterogeneity of subjects*: As the allocation of people to the treatments was based on convenience, the two groups might not have had similar knowledge and backgrounds.
- *Selection*: Participants were selected by the research program leader according to the expected profiles; the PLEvo-Scoping expert, who was not a Fraunhofer employee at that time, had had previous contact with those two participants; the PuLSE-Eco expert was from the same organization as the quasi-experiment participants.

We decided to deal with the lack of statistical tests as proposed by Yin [15], who claims that an effective way of strengthening the results of empirical studies when no statistical test is applicable is to perform them further. Table 1 and its related comments and interpretation (see Subsection 4.5, first paragraph) show that the two groups were comparable indeed, and people with the same role in the different groups had similar profiles (knowledge and background). Therefore, the threat of *Random heterogeneity of subjects* did not appear to be real. Concerning the threat of *Selection*, a question in the profile questionnaire was added that asked about the participant's motivation for taking part in the scoping process of the AAL platform. In addition, the scoping activities were conducted as they would have been conducted with any external customer. Ultimately, this threat did not appear to be real either, because some of the worst evaluations were made by one of the two quasi-experiment participants who had previous contact with the PLEvo-Scoping expert. Regarding the PuLSE-Eco expert, it should be noted that PuLSE-Eco was not the object of study of this quasi-experiment and any possible bias would have affected both treatments.

4.3 Operation

The quasi-experiment took place in the form of one two-day workshop for each treatment. The first part of each workshop was dedicated to the presentation of the application domain, the quasi-experiment's task, as well as relevant information for assuring a common understanding of the PL to be scoped. After that, each member of the scoping team completed the profile questionnaire.

Due to time restrictions, the PLEvo-Scoping expert suggested that the groups divide tasks in some activities according to the participants' role. Group 1 used this approach when identifying and characterizing facts (activity 4, part of the step *Environment Change Anticipation*), when identifying adaptation needs (activity 7, part of the step *Change Impact Analysis*), and when performing the step *PL Evolution Plan* as a whole. Group 2 decided to perform all activities as a group.

Group 2 received extra training and extra time to improve their lists of facts and adaptation needs because the initial number of these was very low (15 and 9, respectively). As scoping the AAL platform PL was a real problem, the goal of this quasi-experiment from the viewpoint of the leader of the AAL research program was to get the highest number and the best quality of results possible from each group. During the analysis of the impact of the adaptation needs on the set of PL features (part of activity 8, in the *Change Impact Analysis* step), the method expert asked Group 1 to define a criterion to distinguish unstable features from stables ones, based on the number of adaptation needs causing changes in the PL features. Group 1 defined that features affected by at least five adaptation needs would be considered unstable. The method expert asked Group 2 to adopt the same criterion.

Another remark concerning the quasi-experiment execution is related to the activities of analyzing the alternative solutions for dealing with relevant adaptation needs and selecting the best alternatives, which are part of the *PL Evolution Plan* step (see Section 2, activities 11 and 12). Due to time restrictions, only the most appropriate alternative solution for each adaptation need was analyzed.

4.4 Data Analysis

Table 1 summarizes the profiles of the quasi-experiment participants. In order to make data summarization and presentation easier, the 5-point ordinal scales were converted into a numerical scale (the higher the number, the better experience, knowledge, or motivation). The domain expert from Group 1 who had the technical viewpoint (column Group 1 - TE) did not answer the question about his capability of providing an overview of the AAL PL and its goals. This has little relevance for the data analysis, because it was not expected that the technical representative would have this capability.

Table 1. Profile of the scoping teams

Profile Item	Group 1			Group 2		
	TE	ME	PLM	TE	ME	PLM
Experience in AAL	3	4	4	2,5	3	4
Knowledge of the AAL platform	5	2	2	5	3	2
Experience in PL Scoping	2	2	1	1	2	4
Technical knowledge in the AAL context	4	3	4	3	4	4
Market knowledge in the AAL context	2	4	4	3	4	4
Capability of providing an overview of the AAL PL and its goals		4	3	2	3	4
Motivation	4	4	4	4	4	5
TE: domain expert with technical viewpoint; ME: domain expert with market viewpoint; PLM: product line manager.						

Table 2 presents a quantitative overview of the scoping process results. The values in parentheses represent the number of facts or adaptation needs that had been given as examples during the extra training and were confirmed by Group 2.

Table 2. Quantitative overview of the workshops

PuLSE-Eco	**Group 1**	**Group 2**
Number of features	52	57
Number of products	7	7
Number of assessed domains	8	7
PLEvo-Scoping	**Group 1**	**Group 2**
Number of facts	30	15 + (11)
Number of adaptation needs	20	12 + (3)
Number of adaptation needs that were planned	8	5

Data Analysis related to Adequacy. Table 3 shows the participants' subjective evaluation of the three levels of support provided by PLEvo-Scoping (see legend at the bottom of Table 3 for details). Each cell gives the number of participants from a certain group who selected the value on the left side to evaluate the support in question. No participant selected the value "Not necessary, nor sufficient" for any level of support, which would be equivalent to 1 on the numeric scale. Participants were asked to indicate missing or annoying information when selecting those values. The annoying (not necessary, but required or provided) information related to Perception of Support 1 (column Support 1 - ST1) was “The relationship between actors’ goals and facts is not so clear. Actors, facts, and adaptation needs would have provided enough support”. The missing (necessary, but not required or provided) information related to Perception of Support 3 (column Support 3 - ST2) was “The basic alternative solutions have to be tailored to the concrete situation, which makes the activity difficult”. Therefore, the values of *# Annoying Information 1* for Group 1 and *# Missing Information 3* for Group 2 were both 1, while the remaining quantitative metrics related to missing and annoying information had the value 0 (see conditions 2 and 3 in Table 5).

Table 3. PLEvo-Scoping support

Selected Value	**Support 1**		**Support 2**		**Support 3**	
	ST1	**ST2**	**ST1**	**ST2**	**ST1**	**ST2**
Missing information (2)						1
Annoying information (3)	1					
Mostly necessary and sufficient (4)	1	2	2	2	1	1
Necessary and sufficient (5)			1	1	2	1

Support 1: support for identifying relevant adaptation needs
Support 2: support for deciding when to introduce an adaptation need and to which products
Support 3: support for analyzing the alternative solutions.

From Table 3, one can calculate, by converting the original ordinal values obtained from the questionnaires into numeric values and applying the arithmetic mean, the values of *Perception of Support 1* (3.5 for Group 1 and 4 for Group 2), *Perception of Support 2* (4.33 for both Group 1 and Group 2), and *Perception of Support 3* (4.67 for Group 1 and 3.67 for Group 2). All of them correspond to the ordinal value of either “Mostly necessary and sufficient” (4) or “Necessary and Sufficient” (5).

Table 4 shows the values obtained for the quantitative metrics. Adaptation needs and technical progress items were not confirmed by the other scoping team if considered irrelevant or outside the scope of the AAL platform PL. Technical progress items are facts related to technology evolution that the PL organization wants to pursue in order to get innovative features to the market as soon as possible.

Table 4. Values of quantitative metrics

Metrics	**Group 1**	**Group 2**
# Added Features	12	3
# Confirmed Added Features	7	2
# Technical Progress Items	8	12
# Confirmed Technical Progress Items	6	6
# Unstable Features	3	8
# Change in Priority due to Volatility	2	0

A few days after the workshop, the quasi-experiment participants were provided with a report on the respective results and asked to provide their perception of PLEvo-Scoping’s adequacy using a 5-point ordinal scale. The ordinal values were converted into numeric ones, and the arithmetic mean was applied. The values obtained for the metric *Perception of Adequacy* were 4.5 (High) according to Group 1, and 4.33

(Medium to High) according to Group 2. No participant selected the values “Low” (1) and “Low to Medium” (2) for the adequacy of the method.

From the defined acceptance criterion (see Table 5) and the above data analysis, proposition P1 (*PLEvo-Scoping is adequate*) was accepted in the context of both groups. While Group 1’s results even satisfied the non-obligatory condition (condition 5), the results of Group 2 did not satisfy it.

Table 5. Acceptance criterion for Adequacy

(1) *Perception of Support 1*, *Perception of Support 2* and *Perception of Support 3* ∈ Positive-Scale1, and
(2) (*# Missing Information 1* + *# Missing Information 2* + *# Missing Information 3*) <= 4 or they can be easily added (1 day of effort), and
(3) (*# Annoying Information 1* + *# Annoying Information 2* + *# Annoying Information 3*) <= 4 or they are already indicated as optional information, and
(4) *Perception of Adequacy* ∈ Positive-Scale2, and
(5) ideally, but not obligatory, *# Changes in Priority due to Volatility* / *# Unstable Features* > 0,50 where

- Positive-Scale1 = {Mostly Necessary and Sufficient, Necessary and Sufficient}, and
- Positive-Scale2 = {Medium to High, High}.

Data Analysis related to Feasibility. The effort for applying both approaches (PuLSE-Eco and PLEvo-Scoping) is described in Table 6, which neither takes into consideration the learning effort nor the scoping experts’ effort. The learning effort related to PLEvo-Scoping was 2.48 person-hours for Group 1 and 2.68 person-hours for Group 2. The scoping experts’ effort represents no loss of information for the quasi-experiment, because the PLEvo-Scoping expert herself did not carry out any activity. The values in parentheses in Table 6 refer to the extra time Group 2 used to improve their lists of facts and adaptation needs, after additional training by the PLEvo-Scoping expert. The PL manager’s effort (column PLM) is also presented factored out of the total effort, because the PL manager from Group 2 managed to take part in more activities of the scoping process than the minimum previously established.

Table 6. Usage effort in person-hours

Phase	**Group 1**	**PLM**	**Group 2**	**PLM**
Product Line Mapping	5.95	1.98	8.60	2.87
Domain Potential Assessment	5.23	0.20	4.02	0.35
Total PuLSE-Eco	11.18	2.18	12.62	3.22
Preparation for Volatility Analysis	0.93		1.50	0.50
Environment Change Anticipation	2.70		3.83 + (0.32)	0.60
Change Impact Analysis	5.25	0.6	6.02 + (0.50)	0.63 + (0.2)
PL Evolution Plan	2.10	1.10	3.90	1.30
Total PLEvo-Scoping	10.98	1.7	15.25 + (0.82)	3.03 + (0.2)
Total Effort (PuLSE + PLEvo)	22.16	3.88	27.87 + (0.82)	6.25 + (0.2)

The quasi-experiment participants evaluated the difficulty of carrying out the PLEvo-Scoping activities. Again, conversion of ordinal values into numeric ones and the arithmetic mean were used. The values obtained for *General Perception of*

Difficulty were 3.27 for Group 1 and 2.83 for Group 2, which represent the ordinal value "Neither Difficult nor Easy".

Furthermore, the quasi-experiment participants were asked to give their perception of the feasibility of PLEvo-Scoping using a 5-point ordinal scale, after being provided with the effort metrics and the method results. The ordinal values were converted into numeric ones, and the arithmetic mean was applied. The values obtained for the metric *Perception of Feasibility* were 4 (Medium to High) according to Group 1, and 3.67 (Medium to High) according to Group 2. No participant selected the values "Low" (1) and "Low to Medium" (2) for the method's feasibility.

Taking into consideration the acceptance criteria defined in Table 7 and the reported data analysis, proposition P2 (*PLEvo-Scoping is feasible*) was accepted in the scope of both scoping teams.

Table 7. Acceptance criterion for Feasibility

(6)	*General Perception of Difficulty* ∈ NDifficult-Scale, and
(7)	*Perception of Feasibility* ∈ Positive-Scale2 where
–	NDifficult-Scale = {Neither Difficult nor Easy, Easy, Very Easy}.

Feedback from the Quasi-Experiment Participants. The quasi-experiment participants were encouraged to register in the questionnaires any comments related to the support provided by PLEvo-Scoping or to the conditions in which the quasi-experiment was performed. The feedback collected was:

- They would like to have 1) more time to perform some activities, 2) a more detailed explanation of PLEvo-Scoping concepts and their relationships ("one slide was not enough"), 3) better tool support for applying the method. PLEvo-Scoping application in this quasi-experiment was supported by an Excel file and the quasi-experiment participants reported it was difficult to switch between different sheets during the execution of an activity.
- They would like to have "a list of adaptation needs that usually apply to all kinds of products or development" (see Section 2, activity 7).
- They reported the difficulty of estimating the business impact and the technical risk of an adaptation need and suggested registering the rationale behind the assignment of values to the attributes (see Section 2, activity 8).
- When elaborating the *PL Evolution Map* (see Section 2, activity 10), it would be useful to have the relationships between the adaptation needs.
- They think "further possibilities to deal with adaptation needs would be reasonable"; alternative solutions are probably missing (see Section 2, activity 11).

Additionally, after having applied PuLSE-Eco only, Group 1 said that it was not easy to think about technical progress, because "it is quite different from thinking about domains, products, and features. We need to switch our minds". PLEvo-Scoping is expected to help them do that.

4.5 Comments and Interpretation of Results

From the values presented in Table 1, we concluded that each participant had the competencies required to perform the role he/she had been assigned to by the AAL research program leader (technical knowledge on the part of the domain expert with the technical viewpoint, market knowledge on the part of the domain expert with the market viewpoint, and the capability of providing an overview of the AAL PL on the part of the PL managers). The values related to *Motivation* were very similar. While Group 1 had higher values for *Experience in AAL*, Group 2 had higher values for *Knowledge of the AAL platform*. Therefore, overall, their domain knowledge can be considered similar as well. The main difference between the two groups is related to *Experience in PL Scoping*, because the PL manager of Group 2 had already participated in a scoping process. As the time spent by the PL managers in the workshops was limited, as PLEvo-Scoping was a new method, and as Group 2 applied the interwoven approach, we do not believe this experience had much influence. Consequently, we considered the two groups to be comparable.

The number of facts and adaptation needs identified by Group 2 (see Subsection 4.3 and Table 2), which initially was too low, and the subsequent extra training and complementary activities to improve it may have been caused by Group 2's choice of performing all activities as a group, not allowing any parallelism. Another factor that may have influenced Group 2's performance is the interweaving of activities from both methods, because the group had to switch their minds between scoping and evolution activities. Despite the differences in Tables 2 and 4, which are justified above, the two approaches for integrating the method into an existing PL scoping process showed similar general results: The method could be applied in just one day, and it was considered adequate, feasible, and neither difficult nor easy to apply.

Concerning the annoying information that was pointed out by a member of Group 1 (see Subsection 4.4 - Data Analysis related to Adequacy), the identification of actors' goals is optional and therefore cannot be considered an annoying request. Furthermore, this information was very useful when Group 2 had identified only 15 facts and 9 adaptation needs and needed some help. The method expert used the actors' goals that had been identified by the group to derive some possible examples of facts and adaptation needs; some of these were considered relevant by the group and accepted. The missing information that was pointed out by a member of Group 2 is really missing information and cannot be added easily. The best way to address it is to build an experience base of alternative solutions for the PL organization.

Contrary to our expectations, the interweaving of activities in treatment 2 did not provide any benefit. We expected that some activities of PuLSE-Eco (especially, the activity *Assess Domains*) would provide insights into some PLEvo-Scoping activities and vice versa, but the workshop format (two consecutive days) and time pressure did not allow the quasi-experiment participants to really benefit from previous activities. The interaction between the two methods must be investigated further.

With regard to the difficulty of applying PLEvo-Scoping, the value for *General Perception of Difficulty* (Neither Difficult nor Easy) is acceptable due to the inherent difficulty of some activities. Furthermore, this result might have been influenced negatively by the short time available for training.

In order to address the feedback obtained from the quasi-experiment participants, we have included in the PLEvo-Scoping activities of characterizing adaptation needs (Section 2, activity 8) and analyzing alternative solutions (Section 2, activity 11) the register of the rationales behind the assignment of values to the attributes. We plan to provide tool support for applying PLEvo-Scoping, which should make the registered and derived relationships between adaptation needs more explicit for the scoping team when elaborating the *PL Evolution Map* (Section 2, activity 10). Concerning the problem of missing alternative solutions, we want to continue to investigate alternative solutions for dealing with adaptation needs together with PL architects and, at the same time, start to collect real occurrences in an experience base of alternative solutions. However, we do not believe it is possible to compile "a list of adaptation needs that usually apply to all kinds of products or development", because adaptation needs are expected to be application-specific and change over time. PLEvo-Scoping provides a list of 35 generic facts instead, so that the scoping team can reflect on whether and how they may apply to the application domain and those facts should lead to the adaptation needs. Moreover, the difficulty of estimating the business impact and the technical risk of an adaptation need cannot be addressed completely, because it is inherent to the problem. Depending on the experience the scoping team has in the domain and with the required technologies, this difficulty is higher or lower.

5 Conclusion

PLEvo-Scoping is a method for supporting PL scoping teams in systematically reasoning about the driving forces of evolution in a certain domain, especially reasoning about who is behind these forces and how their decisions, needs, or achievements may affect the PL infrastructure. Our method allows the PL scoping team to proactively identify and prioritize the adaptation needs that will probably be required in the PL infrastructure and decide about how to deal with them.

The contribution of this paper is the characterization of the adequacy and feasibility of PLEvo-Scoping in practice, meaning according to professionals in charge of scoping an AAL PL. A quasi-experiment was performed to obtain feedback from PL practitioners on how to improve the method and to provide first empirical data on the usage of PLEvo-Scoping, so that other PL organizations can decide on whether to try out the method or not. The method could be applied in just one day, and overall, the quasi-experiment participants perceived it as being adequate and feasible. Those results were really positive, taking into consideration that predicting the future is hard, the quasi-experiment participants applied the method for the first time, the learning effort had to be minimal, and there was no specific tool support.

Although PLEvo-Scoping was applied for just one day, we recommend two to three days for its application, in order to give the PL scoping team enough time to understand the method's underlying concepts and carry out its activities without so much time pressure. This recommendation addresses two of the comments provided by the quasi-experiment participants. In addition, PLEvo-Scoping can also be applied interactively, where the method expert would guide the PL scoping team in performing their activities. The intervention of the PLEvo Scoping expert in this quasi-experiment was kept to a minimum in order not to affect the results.

We think that performing a quasi-experiment is a good means for providing empirical evidence on Product Line engineering technologies, because, compared to case studies and experiment, its intermediate degree of control makes it easier to have PL professionals as subjects while still allowing manipulation of variables and comparison of treatments on some level. In this way, it is possible to convince practitioners as to the applicability of a method in real settings and, at the same time, to provide researchers with scientific evidence on the value of the method.

We intend to perform further empirical studies in order not only to corroborate the results reported in this paper, but also to further analyze the interaction between PLEvo-Scoping and the scoping approach.

Acknowledgments. The authors acknowledge the Alexander von Humboldt Foundation for its financial support and the BelAmi project (BMBF grant number HUN 04/A02) for the opportunity to evaluate PLEvo-Scoping.

References

1. Clements, P., Northrop, L.: Software Product Lines: Practices and Patterns. Addison-Wesley, Reading (2001)
2. Pohl, K., Böckle, G., van der Linden, F.: Software Product Line Engineering: Foundations, Principles, and Techniques. Springer, Heidelberg (2005)
3. Knauber, P., Succi, G.: Perspectives on Software Product Lines. Software Engineering Notes 27(2), 40–45 (2002)
4. Savolainen, J., Kuusela, J.: Volatility Analysis Framework for Product Lines. In: Proc. SSR 2001, Toronto, pp. 133–141 (2001)
5. Bayer, J., Flege, O., Knauber, P., et al.: PuLSE: A Methodology to Develop Software Product Lines. In: Proc. SSR 1999, Los Angeles, pp. 122–131 (1999)
6. Bengtsson, P., Lassing, N., Bosch, J., van Vliet, H.: Analyzing Software Architectures for Modifiability. TR HK-R-RES00/11-SE, University of Karlskrona/Ronneby. Ronneby (2000)
7. Schmid, K.: Planning Software Reuse – A Disciplined Scoping Approach for Software Product Lines. PhD Theses in Experimental Software Engineering. Fraunhofer IRB (2003)
8. John, I., Knodel, J., Lehner, T., et al.: A Practical Guide to Product Line Scoping. In: Proc. SPLC 2006, Baltimore, pp. 3–12 (2006)
9. Villela, K., Dörr, J., Gross, A.: Proactively Managing the Evolution of Embedded System Requirements. In: Proc. RE 2008, Barcelona, pp. 13–22 (2008)
10. John, I., Villela, K., Gross, A.: AAL Platform Product Line – Scoping Results and Recommendations. TR 074.09/E, Fraunhofer IESE, Kaiserslautern (2009) (available upon request)
11. Zelkowitz, M., Wallace, D.: Experimental Models for Validating Technology. IEEE Computer 31(5), 23–31 (1998)
12. Basili, V., Caldiera, G., Rombach, H.: Goal Question Metrics Paradigm. Encyclopedia of Software Engineeering 1, 528–532 (1994)
13. Villela, K., John, I.: Usage of PLEvo-Scoping in the Ambient Assisted Living Domain: A Quasi-Experiment Package. TR 093.09/E, Fraunhofer IESE, Kaiserslautern (2010)
14. Wohlin, C., Runeson, P., Höst, M., et al.: Experimentation in Software Engineering: An Introduction. Kluwer Academic Publishers, Norwell (2000)
15. Yin, R.: Case Study Research: Design and Methods, 3rd edn. Sage Publications, Thousand Oaks (2003)